REMARKS

This application has been reviewed in light of the Office Action dated July 28, 2003. Claims 1, 3-18, 20-35, 37-52, 54-66, 68-80, and 82-94 are presented for examination. Claims 1, 3-6, 18, 20-23, 35, 37-40, 52, 54-58, 66, 68-72, 79, 80, and 82-86 have been amended to define still more clearly what Applicant regards as her invention. Claim 94 has been added to provide Applicant with a more complete scope of protection. Claims 1, 18, 35, 52, 66, 80, and 94 are in independent form. Favorable reconsideration is requested.

Claims 1, 3-10, 18, 20-27, 35, 37-44, 52, 54-61, 66, 68-75, 80, and 82-89 were rejected under 35 U.S.C. §102(b) as being anticipated by IEEE Paper ISBN: 0162-8828; "A Markov Random Field Model-Based Approach to Image Interpretation" (*Modestino et al.*), claims 11, 17, 28, 34, 45, 62-65, 76-79, and 90-93 were rejected under 35 U.S.C. §103(a) as being unpatentable over *Modestino et al.* in view of U.S. Patent No. 5,930,783 (*Li et al.*), claims 12, 29, and 46 were rejected under Section 103(a) as being unpatentable over *Modestino et al.* in view of *Li et al.* and U.S. Patent No. 6,360,234 (*Jain et al.)*, and claims 13-16, 30-33, and 47-51 were rejected under Section 103(a) as being unpatentable over *Modestino et al.* in view of *Jain et al.*

As shown above, Applicant has amended independent claims 1, 18, 35, 52, 66, and 80 in terms that more clearly define the present invention. Applicant submits that these amended independent claims and new independent claim 94, together with the remaining claims dependent thereon, are patentably distinct from the cited prior art for at least the following reasons.

The aspect of the present invention set forth in claim 1 is a method of classifying a digital image. The method segments the digital image into substantially

homogeneous regions and processes the regions to provide a labeled region adjacency graph for the digital image. The labeled region adjacency graph represents spatial adjacency between the regions of the digital image, where at least one of the regions of the labeled region adjacency graph being associated with one of a plurality of predetermined semantic labels. The labeled region adjacency graph is analyzed to identify one or more predetermined patterns of the semantic labels in the labeled region adjacency graph. One or more of a plurality of predetermined stereotypes is assigned to the digital image according to each identified predetermined pattern of the semantic labels in the labeled region adjacency graph. In particular, each of the predetermined stereotypes corresponds to at least one of the predetermined patterns such that the assigned stereotype represents a classification of the digital image based on each predetermined pattern identified in the labeled region adjacency graph.

Among other important features of claim 1 are (1) analyzing the region adjacency graph to identify one or more predetermined patterns of the semantic labels in the labeled region adjacency graph, and (2) assigning one or more of a plurality of predetermined stereotypes to the digital image according to each identified predetermined pattern of the semantic labels in the labeled region adjacency graph, where each of the predetermined stereotypes corresponds to at least one of the predetermined patterns such that the assigned stereotype represents a classification of the digital image based on each predetermined pattern identified in the labeled region adjacency graph.

As described at page 7, lines 1-7, of the specification[1], at each analysis event, the video frame centered on the analysis event is automatically spatially segmented into

---

[1]It is to be understood, of course, that the claim scope is not limited by the details of the described embodiments, which are referred to only to facilitate explanation.

homogeneous regions. These regions and their spatial adjacency properties are represented by a Region Adjacency Graph (RAG). A probabilistic model is then applied to the RAG. The model incorporates feature measurements from the regions of the frame, contextual information from a Region of Interest (ROI) around the frame, and prior knowledge about the various semantic labels that can be associated with the regions of the RAG. Further, as described at page 7, lines 12-15, higher-level expressions, stereotypes, can be used to represent classifications of the video frame or image. Stereotypes can be assigned to video frames, or images, by detecting patterns of region labels and (their) corresponding adjacency in the RAG. This assignment requires a predetermined set of patterns of region labels to be provided, wherein each pattern is associated with a stereotype.

As described at page 7, lines 7-12, the probabilistic model incorporates feature measurements from the regions of the frame, contextual information from a ROI around the frame, and prior knowledge about the various semantic labels that can be associated with the regions of the RAG. These semantic labels (e.g., "person", "sky", "water", "foliage", etc.) are taken from a list which has been typically constructed for an appropriate application domain (e.g., outdoor scenes, weddings, urban scenes, etc).

Many stereotypes can be generated from the analysis of the labeled region adjacency graph (page 14, lines 17-23 of the specification). In one example, (row 1 of Table 1, on page 15 of the specification), a beach scene can be classified as a scene which includes a sky region, a water region and a sand region. That is, the co-occurrence of regions labeled "sky", "water" and "sand", within the labeled RAG, indicate a predetermined pattern of labeled regions which corresponds to a predetermined stereotype (i.e., a beach scene). Further, the stereotype specification, and thus, the associated predetermined pattern can include a requirement such as a minimum size of the regions of the labeled region adjacency

graph (e.g., the water must be larger than some minimum size), adjacency requirements (e.g., the sky must be immediately above the water region) and/or just the existence of a set of labels in the RAG without any further limitations (page 15, rows 2-4 of Table 1 of the specification).

Modestino et al. discusses an image interpretation method based on a Markov Random Field (MRF) model. In the *Modestino et al.* method, an image is segmented into a collection of disjoint regions denoted by *R* that form the nodes of an adjacency graph. Interpretation labels denoted by *I* are assigned to the segmented regions using domain knowledge, extracted feature measurements, and spatial relationships between the various regions, as shown by the example of Fig. 1 of *Modestino et al.*. *Modestino et al.* further discloses that the interpretation labels *I* are modeled as an MRF on the region adjacency graph and that the image interpretation problem is then formulated as a maximum posteriori (MAP) estimation rule, given the domain knowledge and region based measurements.

In the Examiner's response to the Amendment After Final Action dated June 9, 2003, and in making the current rejection under 35 U.S.C. §102(b), the Examiner states that *Modestino et al.* discloses analyzing a region adjacency graph to identify predetermined patterns of the semantic labels, and cites the Abstract, lines 3-10, Figure 1, and pages 607-608, Section II-B of *Modestino et al.* in support thereof.

However, Applicant submits that nothing has been found, or pointed out, in *Modestino et al.* that would teach or suggest analyzing a labeled region adjacency graph to identify one or more predetermined patterns of semantic labels in labeled region adjacency graph. In fact, lines 3-10 of the Abstract merely state that an image is first segmented into a collection of disjoint regions, denoted by *R,* that form the nodes of an adjacency graph. Interpretation labels, denoted by *I,* are assigned to the segmented regions using domain

knowledge, extracted feature measurements, and spatial relationships between the various regions. Lines 10-14 of the Abstract state that the interpretation labels *I* are modeled as an MRF on the region adjacency graph and that the image interpretation problem is then formulated as a maximum posteriori (MAP) estimation rule, given the domain knowledge and region based measurements. In particular, *Modestino et al.* states at lines 1-3 of Section II-B that the image interpretation problem described in Section II-B is restricted to that of labeling the segmented regions, and at lines 9-11 of Section II-B, that a neighborhood system *n* and, consequently, a set of cliques can also be defined on the adjacency graph. As an example, in Fig. 1, and as described at pages 607-608, Section II-B, *Modestino et al.* shows the adjacency graph and all its cliques for a particular synthetic image. *Modestino et al.* then goes on to describe in Section II-B the determination of the vector $I(R)$ representing the object labels assigned to the segmented regions and the optimization of the vector $I(R)$ represented by $Io(R)$.

Accordingly, once the graph has been labeled there is no post-processing or analysis of the region adjacency graph of *Modestino et al.* Nothing has been found in *Modestino et al.* that would teach or suggest analyzing the labeled region adjacency graph, as recited in claim 1.

In making the rejection under 35 U.S.C. §102(b), the Examiner cites the set $L = \{L_1, L_2, \ldots, L_M\}$ disclosed by *Modestino et al.* As the set of all predetermined patterns of regions. This is incorrect. As described at page 607, column 2, of *Modestino et al.*, *L* denotes the set of all interpretation or object (i.e., semantic) labels, *I*. The set $L = \{L_1, L_2, \ldots, L_M\}$ represents a finite set of possible labels that can be used in the labeling process. Accordingly, the set of interpretation labels *L* of *Modestino et al.* merely refers to sets of interpretation labels. For example, for the image of Fig. 1 of *Modestino et al.*, the set of

interpretation labels L = {sky, car, road, field}. Applicant submits that nothing has been found in *Modestino et al.* that would teach or suggest "one or more predetermined patterns of the semantic labels" being identified based on an analysis of "the labeled region adjacency graph," as recited in claim 1.

Applicant further submits that nothing has been found in *Modestino et al.* that would teach or suggest assigning one or more of a plurality of predetermined stereotypes to the digital image according to each identified predetermined pattern of the semantic labels in the labeled region adjacency graph, where each of the predetermined stereotypes corresponds to at least one of the predetermined patterns such that the assigned stereotype represents a classification of the digital image based on each predetermined pattern identified in the labeled region adjacency graph, as recited in claim 1.

As described at page 14, lines 12 to 19, of the specification, many stereotypes can be generated from lower-level analysis of the RAG. In one example, a beach scene can be classified as a scene which includes a sky region, a water region and a sand region. *Modestino et al.* fails to suggest a higher level expression, such as a predetermined stereotype, where each of the predetermined stereotypes corresponds to at least one of the predetermined patterns and each of the predetermined patterns represents a sub-set of the plurality of predetermined semantic labels.

The Examiner equates the perceived "two regions" and "three regions" categories depicted in Table 1 in association with Figs. 1(a) to 1(c) of *Modestino et al.* with stereotypes or higher-level expressions consistent with the definition of stereotype in the specification at page 7, lines 11-17. However, the specification at page 7, lines 16-20, describes that the stereotypes can be assigned to video frames, or images, by detecting patterns of region labels and (their) corresponding adjacency in the RAG. This assignment

requires a predetermined set of patterns of region labels to be provided, where each pattern is associated with a stereotype. Further, as recited in claim 1, one of a plurality of predetermined stereotypes is assigned to the digital image according to each identified predetermined pattern of the semantic labels in the labeled region adjacency graph, where each of the predetermined stereotypes corresponds to at least one of the predetermined patterns. In contrast, the two and three region categories depicted in Table 1 of *Modestino et al.* represent domain knowledge (page 611, column 1, lines 4 to 7). Domain knowledge is applied by *Modestino et al.* during the process of labeling the region adjacency graph, as described with reference to Equation 4, at Section II-B, pages 607 to 608, of *Modestino et al.*

Accordingly, nothing has been found in *Modestino et al.* that would teach or suggest assigning one or more of a plurality of predetermined stereotypes to the digital image according to each identified predetermined pattern of the semantic labels in the labeled region adjacency graph, where each of the predetermined stereotypes corresponds to at least one of the predetermined patterns such that the assigned stereotype represents a classification of the digital image based on each predetermined pattern identified in the labeled region adjacency graph, as recited in claim 1.

For at least the above reasons, Applicant submits that claim 1 is clearly patentable over *Modestino et al.*

Applicant further submits that *Li et al.* fails to remedy the deficiencies of *Modestino et al.* as prior art. *Li et al.* relates to searching and retrieving images contained within a database of images in which both semantic and cognitive methodologies arc utilized. The *Li et al.* method accepts a semantic and cognitive description of an image to be searched from a user, and successively refines the search utilizing semantic and cognitive methodologies and then ranking the results for presentation to the user.

Applicant submits that neither *Li et al.*, *Modestino et al.*, nor any combination thereof (assuming *arguendo* that any such combination would be permissible) teaches or suggests the method as recited in claim 1, of analyzing the labeled region adjacency graph to identify one or more predetermined patterns of the semantic labels in the labeled region adjacency graph. Further, Applicant submits that neither *Li et al.*, *Modestino et al.*, nor any combination thereof, discloses, teaches or suggests assigning one or more of a plurality of predetermined stereotypes the digital image according to each identified predetermined pattern of the semantic labels in the labeled region adjacency graph, wherein each of the predetermined stereotypes corresponds to at least one of the predetermined patterns such that the assigned stereotype represents a classification of the digital image based on each predetermined pattern identified in the labeled region adjacency graph, as recited in claim 1.

Accordingly, Applicant submits that independent claim 1 is clearly patentable over the cited art.

Independent claims 18 and 35 are apparatus and computer program product claims, respectively, corresponding to method claim 1, and are believed to be patentable for at least the same reasons as discussed above in connection with claim 1. Additionally, independent claims 52, 66, 80, and 94 include similar features as those discussed above in connection with claim 1. Accordingly, claims 52, 66, 80, and 94 are believed to be patentable for reasons substantially similar to those discussed above in connection with claim 1.

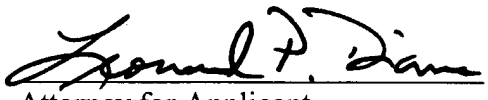
The other rejected claims in this application depend from one or another of the independent claims discussed above, and, therefore, are submitted to be patentable for at least the same reasons. Since each dependent claim is also deemed to define an additional aspect of the invention, individual reconsideration of the patentability of each claim on its own merits is respectfully requested.

A review of the other art of record has failed to reveal anything that, in Applicant' opinion, would remedy the deficiencies of the art discussed above, as applied against the independent claims herein. Therefore, those claims are respectfully submitted to be patentable over the art of record.

In view of the foregoing amendments and remarks, Applicant respectfully requests favorable reconsideration and early passage to issue of the present application.

Applicant's undersigned attorney may be reached in our New York office by telephone at (212) 218-2100. All correspondence should continue to be directed to our below listed address.

Respectfully submitted,



Attorney for Applicant

Registration No. 29,296

FITZPATRICK, CELLA, HARPER & SCINTO
30 Rockefeller Plaza
New York, New York 10112-3801
Facsimile: (212) 218-2200
NYMAIN384632